Plasmalogen biosynthesis by anaerobic bacteria: Identification of a two-gene operon responsible for plasmalogen production in *Clostridium perfringens*

David R. Jackson[1], Chelsi D. Cassilly[1], Damian R. Plichta[2], Hera Vlamakis[2], Hualan Liu[3], Stephen B. Melville[3], Ramnik J. Xavier[2,4,5], Jon Clardy[1]

1. Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston MA 02115.
2. Broad Institute of MIT and Harvard, Cambridge, MA 02142.
3. Department of Biological Sciences, Virginia Tech, Blacksburg, VA, 24061, USA
4. Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114.
5. Center for the Study of Inflammatory Bowel Disease, Massachusetts General Hospital, Boston, MA 02114.

Bacterial Strains	**Description**	**Source**
C. perfringens HN13	Wild-type *C. perfringens* strain	Nariya et. al., 2011 (reference 15)
C. perfringens HN13 Tn library	Previously constructed Tn mutant library	Prof. Stephen Melville
C. perfringens HN13-*ΔPlsA PlsR*	*C. perfringens* HN13 lacking *CPE1194/1195*	This work
E. coli BL21(DE3)	*E. coli* strain for overexpression of pPlsCP	New England Biolabs
E. coli DH5α	*E. coli* strain for cloning	New England Biolabs
E. faecalis OG1RF	*E. faecalis* model strain analyzed for plasmalogen content	A gift from Prof. Danielle Garsin
Plasmids	**Description**	**Source**
pUC19	Plasmid for capturing Tn insertion sites fragments	New England Biolabs
pCM-GalK	Plasmid to create deletions in *C. perfringens* HN13	Nariya et. al., 2011 (reference 15)
pCM-GALK-*Δ1194/1195*	pCM-GALK containing 5' and 3' flanking regions of *CPE1194/CPE1195*	This work
pET-28a	inducible *E. coli* expression vector	Novagen
pPlsCP	pET-28a containing *CPE1194/1195* inserted at NheI/BamHI	This work
pPlsEF	pET-29b containing the *E. faecalis* EF1327 gene inserted at NdeI/XhoI	This work
Primers	**Description**	**Sequence**
OHL21	sequencing primer to map cloned Tn insertion sites	GCAATGAAACACGCCAAAGTAAACAATTTAAGTACCG
OHL22	sequencing primer to map cloned Tn insertion sites	GTTTTATTATTTGGTTGAGTACTTTTTCACTCG
DJ031	forward primer for N-terminal *CPE1194/1195* flanking region with a BamHI site	AAAggatccCAAGTCATCCAGGATGCAAG
DJ032	reverse primer for N-terminal *CPE1194/1195* flanking region	CCATAATTCCTCTAATCATACCCCCTTTTATATTTTCTAGTTC
DJ033	forward primer for C-terminal *CPE1194/1195* flanking region	GAACTAGAAAATATAAAAGGGGGTATGATTAGAGGAATTATGG
DJ034	reverse primer for C-terminal *CPE1194/1195* flanking region with a NheI site	AAAgctagcTTGATACCTTATATCTGTTTTTTCCTAGA
DJ049	forward sequencing primer for *CPE1194/1195* deletion in *C. perfringens* HN13	GGCTTCTAATCCATTTACAGCCTTAAATACAGTGTAC
DJ050	reverse sequencing primer for *CPE1194/1195* deletion in *C. perfringens* HN13	GAGGATTTGATCAAAAGAGCCAAAGAGGG
DJ009	forward primer containing NheI for Pls ORF (*CPE1194/1195*)	CATGgctagcATGTATTACAAAATAGGTATTGATG
DJ010	forward primer containing XhoI for Pls ORF (*CPE1194/1195*)	CTctcgagTCATTCCTGTGCTATAG

Table S1. Bacterial strains, plasmids, and primers used in this study.

		activation				*% identity to:*	
Organism	**Enzyme type**	**Subunit/ domain**	**Mass (kDa)**	**Module**	**NCBI accession**	**PlsA (Act1)**	**PlsA (Act2)**
Clostridium perfringens	plasmalogen synthase (CPE1195 1-288)	PlsA (Act1)	32	activation	WP_025647548	100.0	30.8
Clostridium perfringens	plasmalogen synthase (CPE1195 289-620)	PlsA (Act2)	37	activation	WP_025647548	30.8	100.0
Rhodopseudomonas palustris	benzoyl-CoA reductase	BcrA	48	activation	WP_013503789	22.9	26.0
Rhodopseudomonas palustris	benzoyl-CoA reductase	BcrD	29	activation	WP_013503788	23.4	21.5
Thauera aromatica	benzoyl-CoA reductase	BcrA	48	activation	WP_107220519	22.0	25.2
Thauera aromatica	benzoyl-CoA reductase	BcrD	30	activation	WP_107220520	21.9	22.3
Azoarcus sp. CIB	benzoyl-CoA reductase	BzdP	28	activation	WP_050415405	24.1	23.8
Azoarcus sp. CIB	benzoyl-CoA reductase	BzdQ	33	activation	WP_050415406	23.6	25.1
Ferroglobus placidus	benzoyl-CoA reductase	BzdP	28	activation	WP_012965687	27.1	21.2
Ferroglobus placidus	benzoyl-CoA reductase	BzdQ	33	activation	WP_012965688	27.8	28.7
Fusobacterium nucleatum	(*R*)-2-hydroxyglutaryl-CoA dehydratase	HgdC	27	activation	WP_008694125	25.0	27.6
Acidimonacoccus fermentans	(*R*)-2-hydroxyglutaryl-CoA dehydratase	HgdC	27	activation	WP_012939153	29.8	28.4
Clostridium difficile	(*R*)-2-hydroxyisocaproyl-CoA dehydratase	HadI	28	activation	WP_009888223	26.5	31.0
Clostridium proprionicum	(*R*)-2-lactoyl-CoA dehydratase	LcdC	27	activation	WP_066048110	28.9	28.7
Clostridium sporogenes	(*R*)-phenyllactyl-CoA dehydratase	FldI	28	activation	WP_003492348	27.6	30.0
Archaeoglobus fulgidus	(*R*)-phenyllactyl-CoA dehydratase	HgdC	38	activation	WP_010879451	27.6	30.0
		reduction/					
		dehydratiion				*% identity to:*	
Organism	**Enzyme type**	**Subunit/ domain**	**Mass (kDa)**	**Module**	**NCBI accession**	**PlsA (Red1)**	**PlsR (Red2)**
Clostridium perfringens	plasmalogen synthase (CPE1195 621-975)	PlsA (Red1)	41	reduction/dehydration	WP_025647548	100.0	19.8
Clostridium perfringens	plasmalogen synthase (CPE1194)	PlsR (Red2)	48	reduction/dehydration	WP_025647548	19.8	100.0
Rhodopseudomonas palustris	benzoyl-CoA reductase	BcrC	44	reduction	WP_013503791	19.9	18.3
Rhodopseudomonas palustris	benzoyl-CoA reductase	BcrB	50	reduction	WP_013503790	20.9	23.1
Thauera aromatica	benzoyl-CoA reductase	BcrC	44	reduction	WP_107220517	19.4	20.5
Thauera aromatica	benzoyl-CoA reductase	BcrB	49	reduction	WP_107220518	19.8	21.5
Azoarcus sp. CIB	benzoyl-CoA reductase	BzdN	43	reduction	WP_050415404	21.6	18.9
Azoarcus sp. CIB	benzoyl-CoA reductase	BzdO	51	reduction	WP_050418279	17.7	17.3
Ferroglobus placidus	benzoyl-CoA reductase	BzdN	45	reduction	WP_148212195	21.0	20.1
Ferroglobus placidus	benzoyl-CoA reductase	BzdO	49	reduction	WP_048086821	20.1	19.3
Fusobacterium nucleatum	(*R*)-2-hydroxyglutaryl-CoA dehydratase	HgdA	50	dehydration	WP_005909813	19.3	19.9
Fusobacterium nucleatum	(*R*)-2-hydroxyglutaryl-CoA dehydratase	HgdB	42	dehydration	WP_005909812	22.9	21.5
Acidimonacoccus fermentans	(*R*)-2-hydroxyglutaryl-CoA dehydratase	HgdA	54	dehydration	WP_012939152	21.0	18.0
Acidimonacoccus fermentans	(*R*)-2-hydroxyglutaryl-CoA dehydratase	HgdB	42	dehydration	WP_012939151	21.5	20.1
Clostridium difficile	(*R*)-2-hydroxyisocaproyl-CoA dehydratase	HadB	46	dehydration	WP_009888224	22.0	21.9
Clostridium difficile	(*R*)-2-hydroxyisocaproyl-CoA dehydratase	HadC	42	dehydration	WP_009888224	20.0	22.7
Clostridium proprionicum	(*R*)-2-lactoyl-CoA dehydratase	LcdA	47	dehydration	WP_066048111	19.2	22.1
Clostridium proprionicum	(*R*)-2-lactoyl-CoA dehydratase	LcdB	42	dehydration	WP_066048114	19.4	23.5
Clostridium sporogenes	(*R*)-phenyllactyl-CoA dehydratase	FldB	46	dehydration	WP_003492352	23.5	20.8
Clostridium sporogenes	(*R*)-phenyllactyl-CoA dehydratase	FldC	43	dehydration	WP_003492354	20.0	24.3
Archaeoglobus fulgidus	(*R*)-phenyllactyl-CoA dehydratase	HgdA	45	dehydration	WP_010879450	21.2	22.3
Archaeoglobus fulgidus	(*R*)-phenyllactyl-CoA dehydratase	HgdB	27	dehydration	WP_010879449	20.6	23.5

Table S2: Comparison of PlsA domains and PlsR to known HAD/BCR proteins.

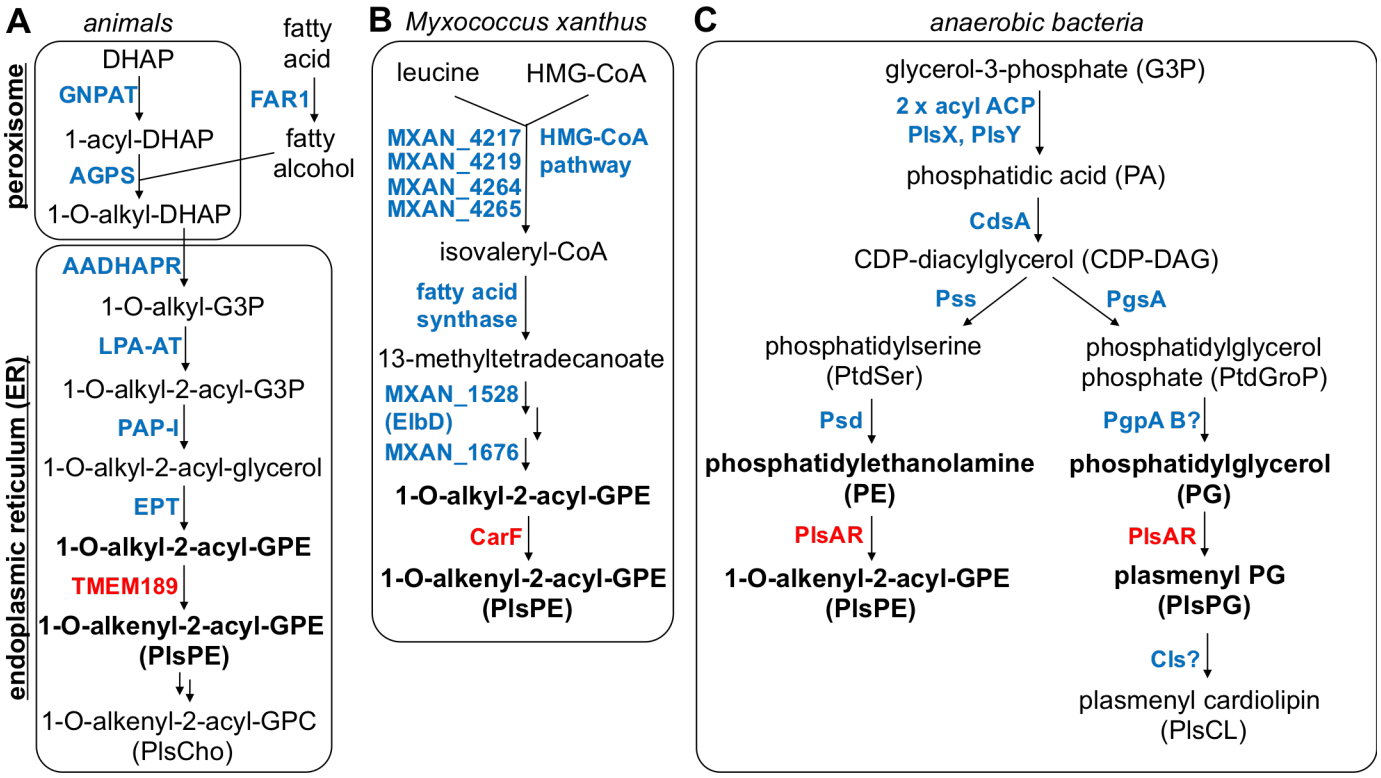

Figure S1. A) Plasmalogen biosynthesis in animals. B) Plasmalogen biosynthesis in the aerobic bacterium *Myxococcus xanthus*. C) Plasmalogen biosynthesis in anaerobes (proposed).

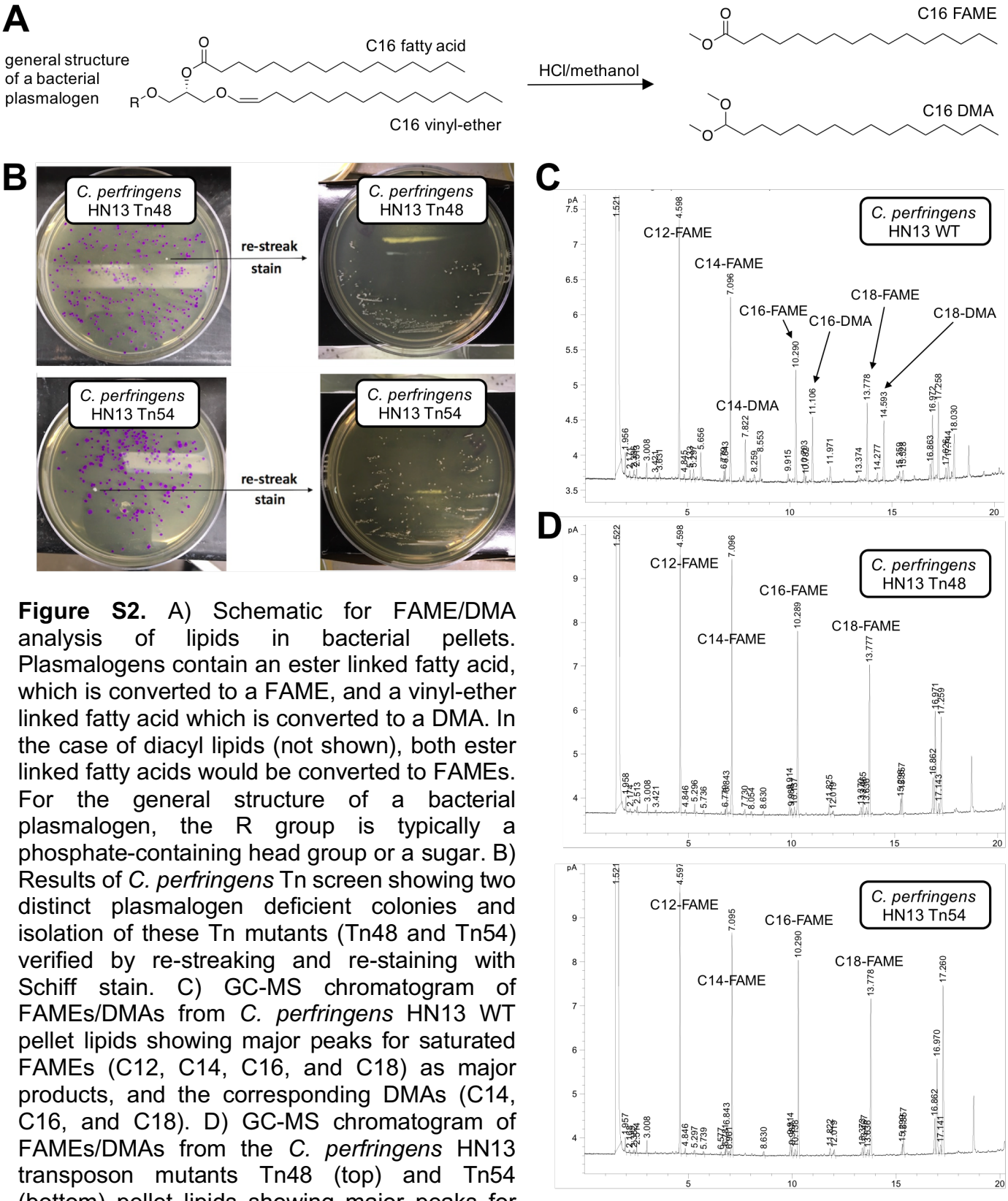

Figure S2. A) Schematic for FAME/DMA analysis of lipids in bacterial pellets. Plasmalogens contain an ester linked fatty acid, which is converted to a FAME, and a vinyl-ether linked fatty acid which is converted to a DMA. In the case of diacyl lipids (not shown), both ester linked fatty acids would be converted to FAMEs. For the general structure of a bacterial plasmalogen, the R group is typically a phosphate-containing head group or a sugar. B) Results of *C. perfringens* Tn screen showing two distinct plasmalogen deficient colonies and isolation of these Tn mutants (Tn48 and Tn54) verified by re-streaking and re-staining with Schiff stain. C) GC-MS chromatogram of FAMEs/DMAs from *C. perfringens* HN13 WT pellet lipids showing major peaks for saturated FAMEs (C12, C14, C16, and C18) as major products, and the corresponding DMAs (C14, C16, and C18). D) GC-MS chromatogram of FAMEs/DMAs from the *C. perfringens* HN13 transposon mutants Tn48 (top) and Tn54 (bottom) pellet lipids showing major peaks for saturated FAMEs (C12, C14, C16, and C18) as major products, and the absence of any corresponding DMAs.

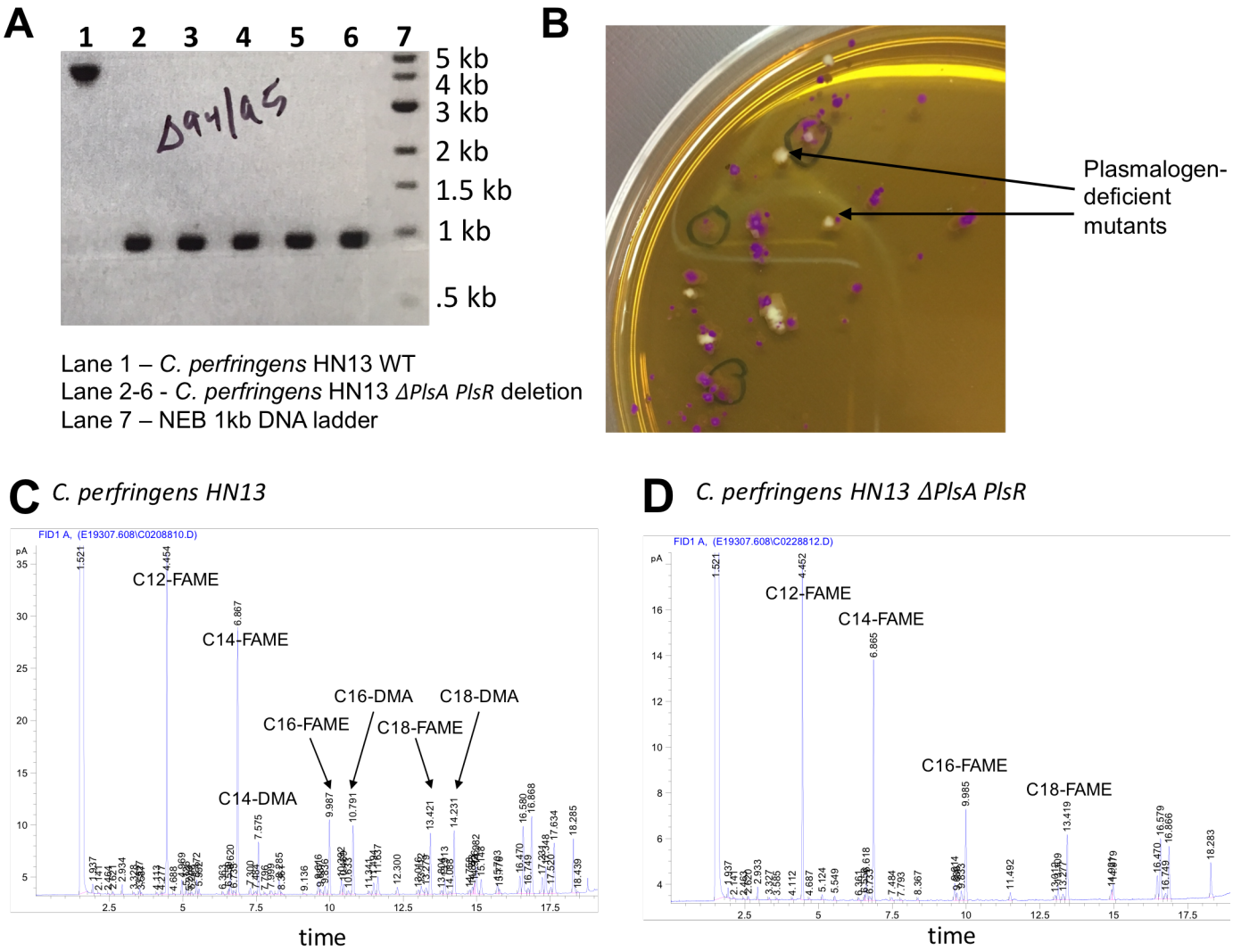

Figure S3. A) PCR verification of *CPE1194/1195 (ΔPlsA PlsR)* deletion. B) Isolation of *C. perfringens* HN13 *ΔPlsA PlsR* deletion mutants by Schiff staining. C) GC-MS chromatogram of FAMEs/DMAs from WT *C. perfringens* HN13 pellet lipids showing major peaks for saturated FAMEs (C12, C14, C16, and C18) as major products, and the corresponding DMAs (C14, C16, and C18). D) GC-MS chromatogram of FAMEs/DMAs from the *C. perfringens* HN13 *ΔPlsA PlsR* deletion mutant pellet lipids showing major peaks for saturated FAMEs (C12, C14, C16, and C18) as major products, and the absence of any corresponding DMAs.

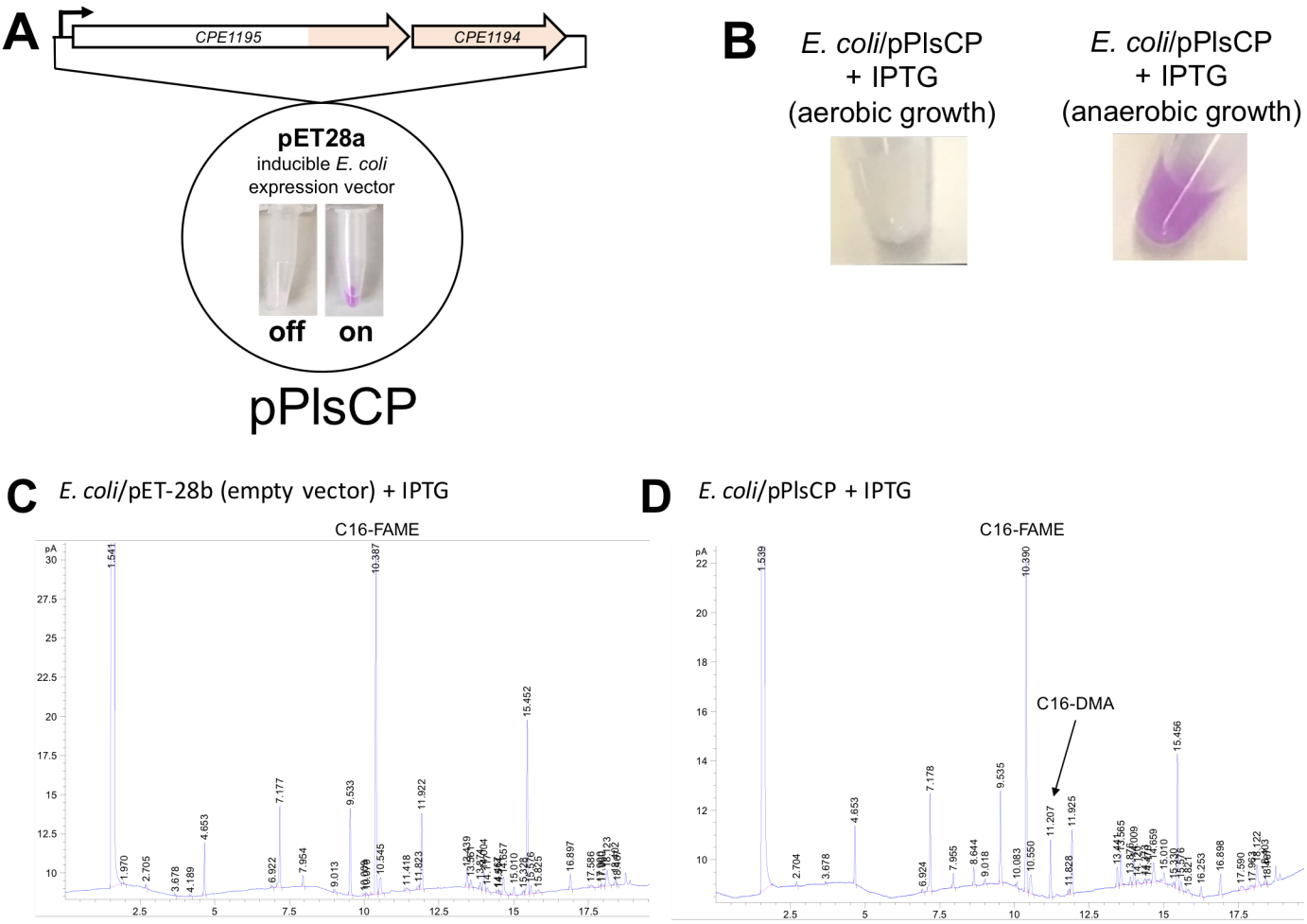

Figure S4. A) A schematic representation the pET28a-based expression vector for *C. perfringens* CPE1194/1195 operon. B) Pellets from 5 ml cultures *E. coli* BL21(DE3) harboring pPlsCP with IPTG induction grown aerobically (top) and anaerobically (bottom) after staining with Schiff. C) GC-MS analysis of *E. coli* BL21(DE3) harboring pET28a (empty) showing only FAMEs and no DMAs, the most abundant peak is saturated C16-FAME. D) GC-MS analysis of *E. coli* BL21(DE3) harboring pPlsCP showing similar FAMEs and an additional peak corresponding to C16-DMA.

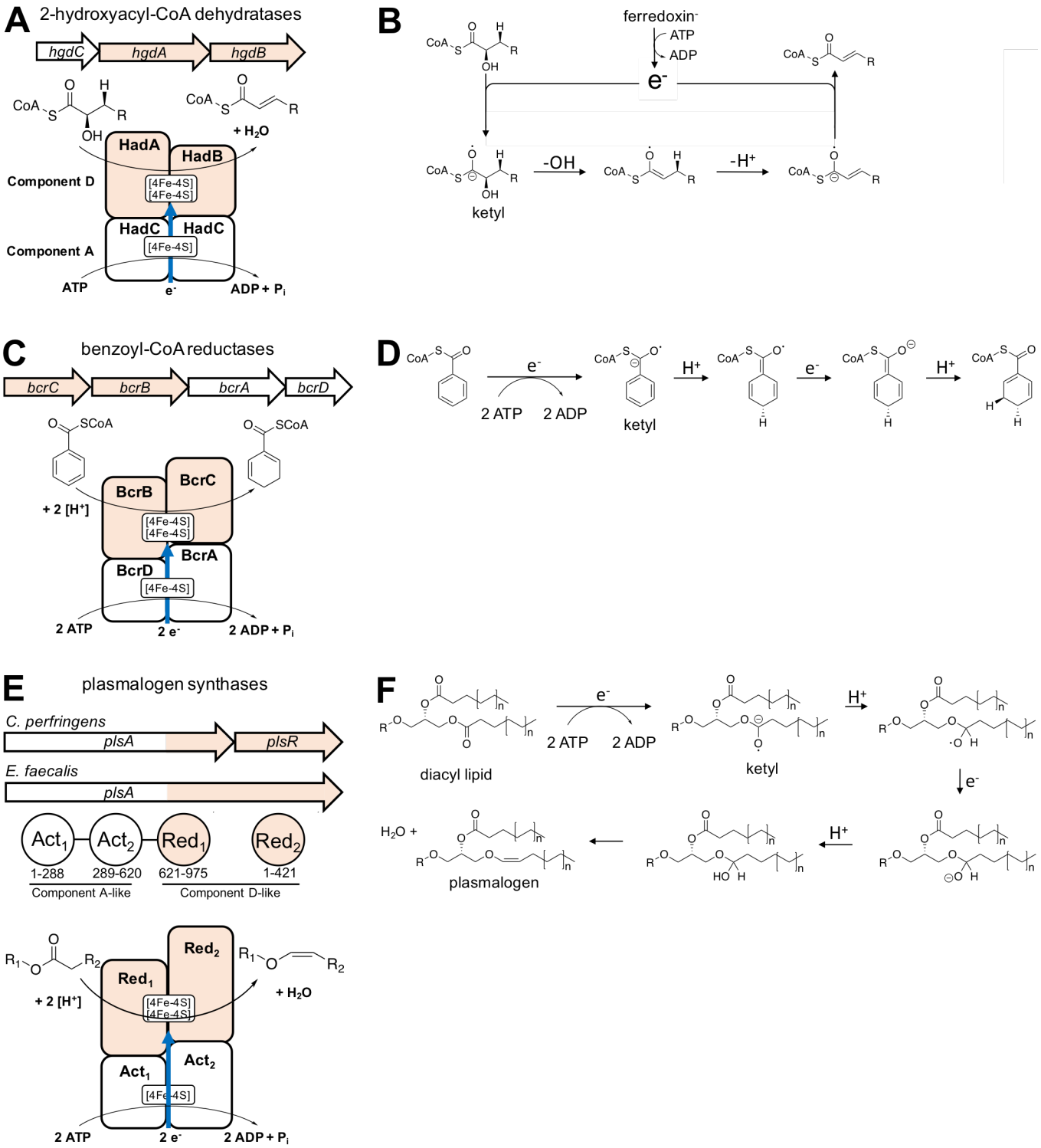

Figure S5. A) A typical HAD gene cluster and proposed enzyme architecture. B) HAD reaction mechanism. C) A typical BCR gene cluster and proposed enzyme architecture. D) BCR reaction mechanism. E) Different genetic architectures of the *pls* operon and proposed enzyme architecture. F) The plasmalogen synthase reaction mechanism based on the current BCR model involving two sequential single electron reductions, followed by protonation and elimination of water to yield a cis enol ether.

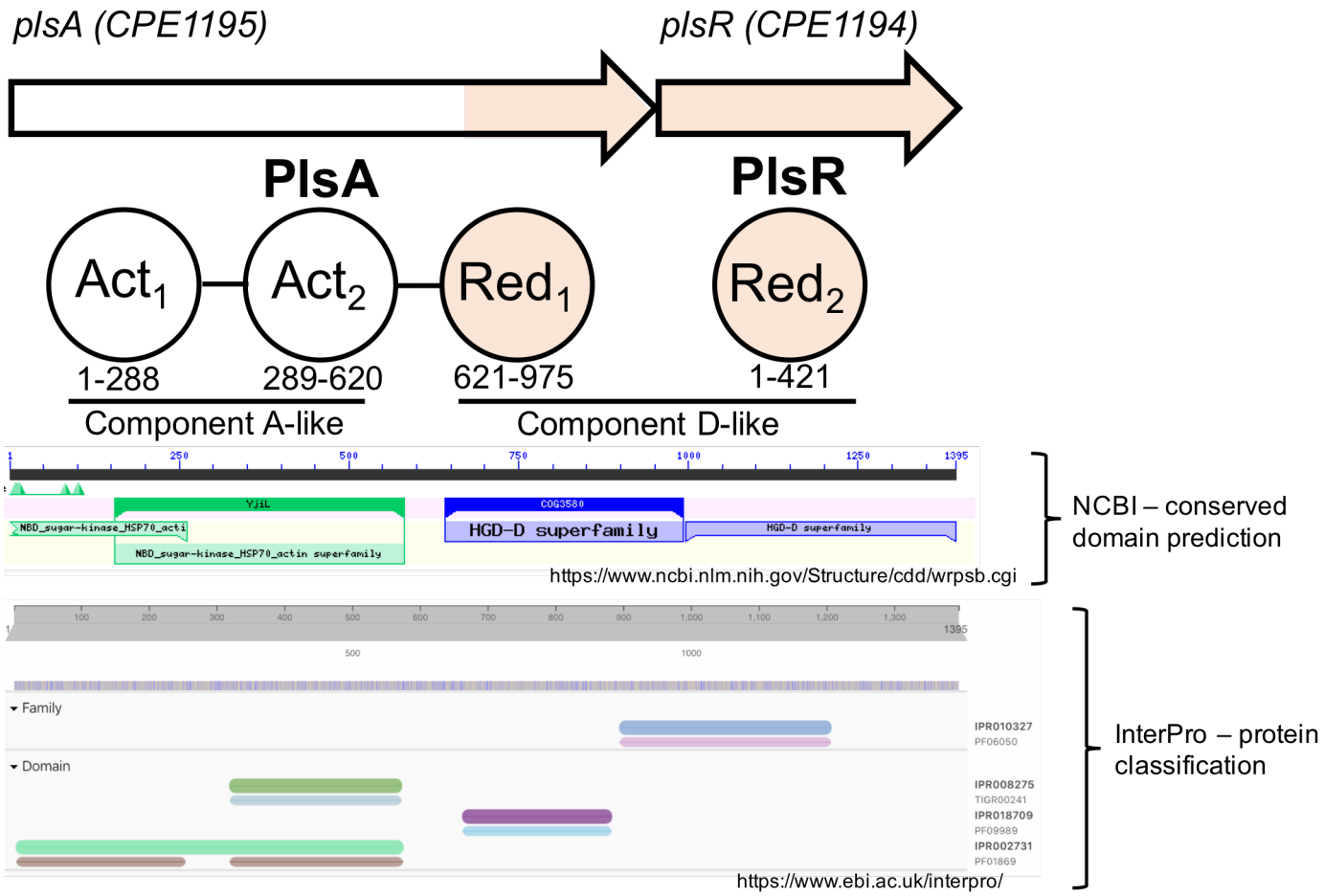

Figure S6. Predicted domain analysis of *CPE1194/1195* using NCBI and InterPro. *CPE1195* (*plsA*) is composed of two CoA-activase domains (Act_1 and Act_2, component A) and one reduction/dehydration domain (Red_1, component D). *CPE1194* (*plsR*) is composed of a single reduction/dehydration domain (Red_2, component D).

A

	1	2	3	4	5	6
1: A_fulgidus_HgdC	100.00	49.00	49.80	45.42	46.22	45.82
2: F_nucleatum_HgdC	49.00	100.00	63.32	54.41	57.20	54.83
3: A_fermentas_HgdC	49.80	63.32	100.00	53.70	56.76	57.98
4: C_difficile_HadI	45.42	54.41	53.70	100.00	57.85	55.98
5: C_sporogenes_FldI	46.22	57.20	56.76	57.85	100.00	55.98
6: C_proprionicum_LcdC	45.82	54.83	57.98	55.98	55.98	100.00

B

	1	2	3	4	5	6	7	8
1: R_palustris_BcrD	100.00	71.12	25.00	23.88	30.12	28.24	30.92	28.46
2: T_aromatica_BcrD	71.12	100.00	27.99	25.00	32.05	29.77	32.06	26.48
3: R_palustris_BcrA	25.00	27.99	100.00	69.50	33.68	31.47	24.25	23.85
4: T_aromatica_BcrA	23.88	25.00	69.50	100.00	33.90	32.17	25.09	26.64
5: Azoarcus_CIB_BzdQ	30.12	32.05	33.68	33.90	100.00	60.07	30.74	28.63
6: F_placidus_BzdQ	28.24	29.77	31.47	32.17	60.07	100.00	37.21	32.81
7: Azoarcus_CIB_BzdP	30.92	32.06	24.25	25.09	30.74	37.21	100.00	43.02
8: F_placidus_BzdP	28.46	26.48	23.85	26.64	28.63	32.81	43.02	100.00

Figure S7. A) Sequence identify matrix of HAD activation domains. B) Sequence identify matrix of BCR activation domains.

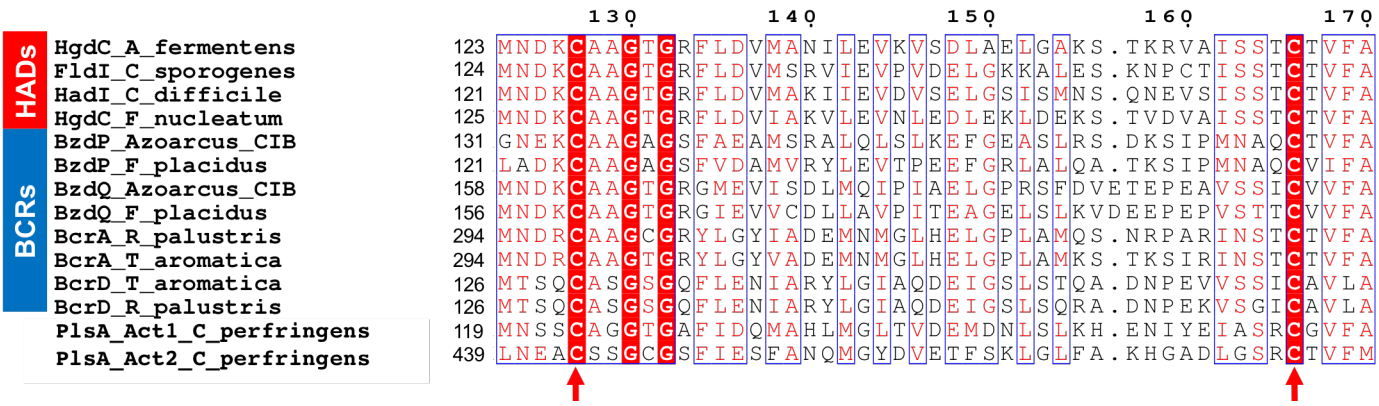

Figure S8. Protein sequence alignment of activation domains from HADs, BCRs, and PlsA (Act_1, AAs 1-288) and PlsA (Act_2, AAs 289-620) showing conserved cysteines for iron-sulfur cluster ligation.

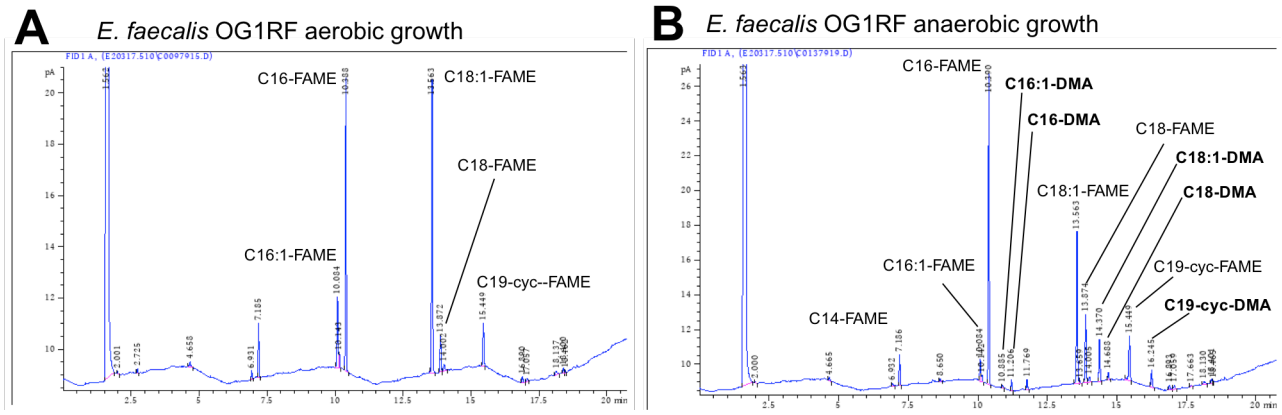

Figure S9. A) GC-MS analysis of *E. faecalis* grown aerobically showing only FAMEs present. B) GC-MS analysis of *E. faecalis* grown anaerobically showing FAMEs and plasmalogen-derived DMAs.

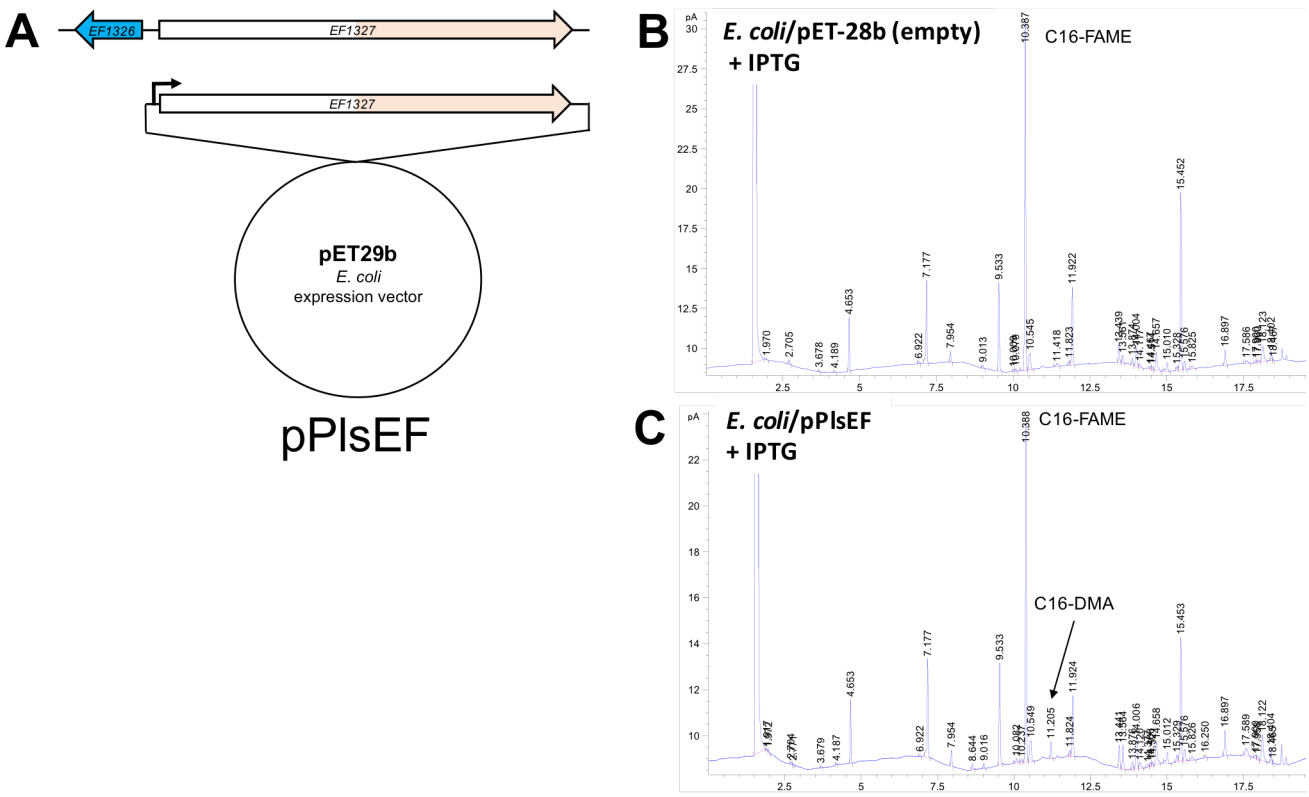

Figure S10: A) *E. coli* expression plasmid for the single gene *pls* operon from *E. faecalis* OG1RF. B) GC-MS analysis of *E. coli* BL21(DE3) harboring pET28a (empty) showing only FAMEs and no DMAs, the most abundant peak is saturated C16-FAME. C) GC-MS analysis of *E. coli* BL21(DE3) harboring pPlsEF showing similar FAMEs and an additional peak corresponding to C16-DMA.